GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: February 26, 2002, 13:21:28 ; Search time 205.96 Seconds
(without alignments)
2834.718 Million cell updates/sec

Title: US-09-602-833A-3
Perfect score: 681
Sequence: 1 atgagaattctggatctgcc.....cttttagccttcaactttga 681

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 930621 seqs, 428662619 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_1101:*
1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	681	100.0	681	22	AAF24903	Nucleotide sequenc
	2	681	100.0	1116	22	AAF24902	Nucleotide sequenc
	3	129	18.9	2056	22	AAH17218	Human cDNA sequenc
	4	19	2.8	327	21	AAC00689	Human secreted pro
	5	19	2.8	413	21	AAC06478	Human secreted pro
	6	19	2.8	1185	16	AAT03478	Transcription fact
	7	19	2.8	1816	21	AAC78156	Human cancer assoc
	8	19	2.8	3072	21	AAA75580	DNA encoding a mou
	9	18	2.6	811	19	AAX13947	H. pylori GHPO 127
c	10	18	2.6	2324	22	AAD10125	Mouse serotransfer
c	11	18	2.6	4118	17	AAT44520	NTHI hxuC + hxuB g
	12	18	2.6	6253	20	AAX13097	Enterococcus faeca
c	13	18	2.6	1038602	20	AAZ01425	Complete genome se
c	14	17	2.5	156	22	AAI19482	Probe #9415 for ge
c	15	17	2.5	156	22	AAI20472	Probe #10405 for g
c	16	17	2.5	156	22	AAI44677	Probe #13363 used
c	17	17	2.5	156	22	AAI45679	Probe #14365 used
c	18	17	2.5	156	22	AAI05210	Probe #5201 used t
c	19	17	2.5	156	22	AAI06170	Probe #6161 used t
	20	17	2.5	296	18	AAV78820	Staphylococcus aur
c	21	17	2.5	390	21	AAA67203	Pinus radiata alph
c	22	17	2.5	418	21	AAA67201	Pinus radiata alph
c	23	17	2.5	441	21	AAA67198	Pinus radiata alph
	24	17	2.5	443	22	AAH35681	Human colon cancer
	25	17	2.5	460	22	AAI15848	Probe #5781 for ge
	26	17	2.5	460	22	AAI37733	Probe #6419 used t
c	27	17	2.5	479	22	AAI10199	Probe #132 for gen
c	28	17	2.5	479	22	AAI31451	Probe #137 used to
c	29	17	2.5	479	22	AAI00144	Probe #135 used to
c	30	17	2.5	481	21	AAA67199	Pinus radiata alph
c	31	17	2.5	488	22	AAI11251	Probe #1184 for ge
c	32	17	2.5	488	22	AAI32517	Probe #1203 used t
c	33	17	2.5	488	22	AAI01166	Probe #1157 used t
	34	17	2.5	955	20	AAX57430	Rat U3 gene trap d
	35	17	2.5	1163	21	AAC39227	Arabidopsis thalia
c	36	17	2.5	1299	21	AAF07517	Fusarium venenatum
c	37	17	2.5	1309	22	AAI61059	Human polynucleoti
	38	17	2.5	1331	22	AAI59273	Human polynucleoti
c	39	17	2.5	1355	9	AAN82025	Fragment containin
c	40	17	2.5	1597	12	AAQ10867	T18 oncogene. Mus
c	41	17	2.5	1597	12	AAQ14048	Human OT18 clone p
	42	17	2.5	1652	21	AAC96935	Human secreted pro
c	43	17	2.5	2127	22	AAH13909	Human cDNA sequenc
c	44	17	2.5	2197	22	AAH18389	Human cDNA sequenc
	45	17	2.5	2259	22	AAH52477	S. epidermidis ope

ALIGNMENTS

RESULT 1
AAF24903
ID AAF24903 standard; cDNA; 681 BP.
XX
AC AAF24903;
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of a human SGT4-2 polypeptide.
XX
KW Human; SGT4; signal transduction; guanosine triphosphate binding protein;
KW GTP binding protein; cancer; immune response; nutritional source;
KW animal feed; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..681
FT /*tag= a
FT /product= "SGT4"
XX
PN WO200078959-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-US17248.
XX
PR 23-JUN-1999; 99US-0140627.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Turner AC, Zambrowicz B, Nehls M, Friedrich GA, Sands AT;
XX
DR WPI; 2001-032329/04.

```
DR   P-PSDB; AAB31564.
XX
PT   New SGT4 genes and proteins, useful for diagnosing and treating
PT   disorders involving inappropriate regulation of a signal transduction
PT   mechanism e.g. cancer -
XX
PS   Claim 1; Fig 3; 82pp; English.
XX
CC   The present sequence encodes a human SGT4 polypeptide. SGT4 polypeptides
CC   are involved in signal transduction pathways regulated by guanosine
CC   triphosphate (GTP) binding proteins). SGT4 polynucleotides and
CC   polypeptides are for diagnosing and treating conditions related to a
CC   signal transduction mechanism involving SGT4 such as cancer. In
CC   addition, it can be used to detect the expression of SGT4 as markers of
CC   specific cells and tissues such as neuronal tissue, heart, liver,
CC   pancreas and adrenal gland. They are also useful for the construction of
CC   transgenic and knockout animals for studying SGT4 function in vivo and
CC   for the screening of SGT4 (ant)agonists in an animal model. Other more
CC   general uses include: as molecular weight markers on Southern gels; as
CC   chromosome markers or tags; as probes; for selecting and making
CC   oligomers for attachment to a gene chip; to raise anti-protein or
CC   anti-DNA antibodies or to elicit immune response. They are also
CC   also be used as nutritional sources or supplements such as in animal
CC   feed.
XX
SQ   Sequence 681 BP; 212 A; 138 C; 142 G; 189 T; 0 other;


  Query Match             100.0%;  Score 681;  DB 22;  Length 681;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  681;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 atgagaattctggatctgccaaaaaaccaaatctcacatcttccagcagaaatcggttgt 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 atgagaattctggatctgccaaaaaaccaaatctcacatcttccagcagaaatcggttgt 60

Qy      61 ttgaagaacctgaaagaactcaatgtgggtttcaactatctgaagagcattcctccagaa 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ttgaagaacctgaaagaactcaatgtgggtttcaactatctgaagagcattcctccagaa 120

Qy     121 ttgggagattgtgaaaatctagagagactggattgttctggaaatctagaattaatggag 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ttgggagattgtgaaaatctagagagactggattgttctggaaatctagaattaatggag 180

Qy     181 ctgccctttgaattaagtaatttgaagcaagttacatttgtagatatctcagcaaacaag 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ctgccctttgaattaagtaatttgaagcaagttacatttgtagatatctcagcaaacaag 240

Qy     241 ttttccagtgtcccaatctgtgtcctgcggatgtcgaatttgcagtggttggatatcagc 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ttttccagtgtcccaatctgtgtcctgcggatgtcgaatttgcagtggttggatatcagc 300

Qy     301 agcaataacctgaccgacctgccgcaagatatagacaggctagaggagctgcagagcttt 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 agcaataacctgaccgacctgccgcaagatatagacaggctagaggagctgcagagcttt 360

Qy     361 ctcttgtataaaaacaagttgacctaccttccctattccatgctgaacctgaagaagctc 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ctcttgtataaaaacaagttgacctaccttccctattccatgctgaacctgaagaagctc 420

Qy     421 actctgttagtcgtcagtggggaccatttggtggagctcccaactgccctttgtgactca 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 actctgttagtcgtcagtggggaccatttggtggagctcccaactgccctttgtgactca 480

Qy     481 tccacacctttaaaatttgtaagccttatggacaatcctattgataatgcccaatgtgaa 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 tccacacctttaaaatttgtaagccttatggacaatcctattgataatgcccaatgtgaa 540

Qy     541 gatggcaatgaaataatggaaagtgaacgggatcgccaacattttgataaagaagttatg 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 gatggcaatgaaataatggaaagtgaacgggatcgccaacattttgataaagaagttatg 600

Qy     601 aaagcctatattgaagaccttaaagaaagagaatctgttcccagctataccaccaaagtg 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 aaagcctatattgaagaccttaaagaaagagaatctgttcccagctataccaccaaagtg 660

Qy     661 tcttttagccttcaactttga 681
           |||||||||||||||||||||
Db     661 tcttttagccttcaactttga 681


RESULT   2
AAF24902
ID   AAF24902 standard; cDNA; 1116 BP.
XX
AC   AAF24902;
XX
DT   20-APR-2001  (first entry)
XX
DE   Nucleotide sequence of a human SGT4-1 polypeptide.
XX
KW   Human; SGT4; signal transduction; guanosine triphosphate binding protein;
KW   GTP binding protein; cancer; immune response; nutritional source;
KW   animal feed; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1116
FT                   /*tag= a
FT                   /product= "SGT4"
XX
PN   WO200078959-A1.
XX
PD   28-DEC-2000.
XX
PF   22-JUN-2000; 2000WO-US17248.
XX
PR   23-JUN-1999;   99US-0140627.
XX
PA   (LEXI-) LEXICON GENETICS INC.
XX
PI   Turner AC,  Zambrowicz B,  Nehls M,  Friedrich GA,  Sands AT;
XX
DR   WPI; 2001-032329/04.
DR   P-PSDB; AAB31563.
XX
PT   New SGT4 genes and proteins, useful for diagnosing and treating
PT   disorders involving inappropriate regulation of a signal transduction
PT   mechanism e.g. cancer -
XX
PS   Claim 1; Fig 1; 82pp; English.
XX
CC   The present sequence encodes a human SGT4 polypeptide. SGT4 polypeptides
CC   are involved in signal transduction pathways regulated by guanosine
CC   triphosphate (GTP) binding proteins). SGT4 polynucleotides and
CC   polypeptides are for diagnosing and treating conditions related to a
CC   signal transduction mechanism involving SGT4 such as cancer. In
CC   addition, it can be used to detect the expression of SGT4 as markers of
CC   specific cells and tissues such as neuronal tissue, heart, liver,
CC   pancreas and adrenal gland. They are also useful for the construction of
CC   transgenic and knockout animals for studying SGT4 function in vivo and
CC   for the screening of SGT4 (ant)agonists in an animal model. Other more
CC   general uses include: as molecular weight markers on Southern gels; as
CC   chromosome markers or tags; as probes; for selecting and making
CC   oligomers for attachment to a gene chip; to raise anti-protein or
CC   anti-DNA antibodies or to elicit immune response. They are also
CC   also be used as nutritional sources or supplements such as in animal
CC   feed.
XX
SQ   Sequence 1116 BP; 343 A; 224 C; 265 G; 284 T; 0 other;


  Query Match             100.0%;  Score 681;  DB 22;  Length 1116;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  681;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 atgagaattctggatctgccaaaaaaccaaaatctcacatcttccagcagaaatcggttgt 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    436 atgagaattctggatctgccaaaaaaccaaaatctcacatcttccagcagaaatcggttgt 495

Qy     61 ttgaagaacctgaaagaactcaatgtgggtttcaactatctgaagagcattcctccagaa 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    496 ttgaagaacctgaaagaactcaatgtgggtttcaactatctgaagagcattcctccagaa 555

Qy    121 ttgggagattgtgaaaatctagagagactggattgttctggaaatctagaattaatggag 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    556 ttgggagattgtgaaaatctagagagactggattgttctggaaatctagaattaatggag 615

Qy    181 ctgccctttgaattaagtaatttgaagcaagttacatttgtagatatctcagcaaacaag 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    616 ctgccctttgaattaagtaatttgaagcaagttacatttgtagatatctcagcaaacaag 675

Qy    241 ttttccagtgtcccaatctgtgtcctgcggatgtcgaatttgcagtggttggatatcagc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    676 ttttccagtgtcccaatctgtgtcctgcggatgtcgaatttgcagtggttggatatcagc 735

Qy    301 agcaataacctgaccgacctgccgcaagatatagacaggctagaggagctgcagagcttt 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    736 agcaataacctgaccgacctgccgcaagatatagacaggctagaggagctgcagagcttt 795

Qy    361 ctcttgtataaaaacaagttgacctaccttccctattccatgctgaacctgaagaagctc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    796 ctcttgtataaaaacaagttgacctaccttccctattccatgctgaacctgaagaagctc 855

Qy    421 actctgttagtcgtcagtggggaccatttggtggagctcccaactgccctttgtgactca 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    856 actctgttagtcgtcagtggggaccatttggtggagctcccaactgccctttgtgactca 915

Qy    481 tccacacctttaaaatttgtaagccttatggacaatcctattgataatgcccaatgtgaa 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    916 tccacacctttaaaatttgtaagccttatggacaatcctattgataatgcccaatgtgaa 975

Qy    541 gatggcaatgaaataatggaaagtgaacgggatcgccaacattttgataaagaagttatg 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    976 gatggcaatgaaataatggaaagtgaacgggatcgccaacattttgataaagaagttatg 1035

Qy    601 aaagcctatattgaagaccttaaagaaagagaatctgttcccagctataccaccaaagtg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1036 aaagcctatattgaagaccttaaagaaagagaatctgttcccagctataccaccaaagtg 1095

Qy    661 tcttttagccttcaactttga 681
          |||||||||||||||||||||
Db   1096 tcttttagccttcaactttga 1116
```

```
RESULT   3
AAH17218
ID   AAH17218 standard; cDNA; 2056 BP.
XX
AC   AAH17218;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:16594.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 8; SEQ ID 16594; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 2056 BP; 642 A; 394 C; 495 G; 525 T; 0 other;
```

```
  Query Match            18.9%;  Score 129;  DB 22;  Length 2056;
  Best Local Similarity  100.0%;  Pred. No. 4.2e-55;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    553 ataatggaaagtgaacgggatcgccaacattttgataaagaagttatgaaagcctatatt 612
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1294 ataatggaaagtgaacgggatcgccaacattttgataaagaagttatgaaagcctatatt 1353

Qy    613 gaagaccttaaagaaagagaatctgttcccagctataccaccaaagtgtcttttagcctt 672
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1354 gaagaccttaaagaaagagaatctgttcccagctataccaccaaagtgtcttttagcctt 1413

Qy    673 caactttga 681
          |||||||||
Db   1414 caactttga 1422
```

```
RESULT   4
AAC00689
ID   AAC00689 standard; cDNA; 327 BP.
XX
AC   AAC00689;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 687.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
```

OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
DR P-PSDB; AAG00683.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 687; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 327 BP; 121 A; 48 C; 85 G; 73 T; 0 other;

Query Match 2.8%; Score 19; DB 21; Length 327;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 tgaagaacctgaaagaact 80
||||||||||||||||||
Db 88 tgaagaacctgaaagaact 106

RESULT 5
AAC06478
ID AAC06478 standard; cDNA; 413 BP.
XX
AC AAC06478;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 10553.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 10553; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 413 BP; 138 A; 76 C; 103 G; 95 T; 1 other;

Query Match 2.8%; Score 19; DB 21; Length 413;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 tgaagaacctgaaagaact 80
||||||||||||||||||
Db 173 tgaagaacctgaaagaact 191

RESULT 6
AAT03478
ID AAT03478 standard; DNA; 1185 BP.
XX
AC AAT03478;
XX
DT 06-JUN-1996 (first entry)
XX
DE Transcription factor BTF2 complex p44 subunit gene.
XX
KW Transcription factor; BTF2; subunit; kinase; ATPase; helicase; HeLa; PCR;
KW reconstitution; in vitro transcription system; probe; primer; antibody;
KW amplification; microsequence; cancer; skin melanoma; xeroderma; UV light;
KW Cockayne syndrome; skin pigmentation disorder; sensitivity; ss.
XX
OS Homo sapiens.
XX
PN WO9529245-A2.
XX
PD 02-NOV-1995.
XX
PF 25-APR-1995; 95WO-FR00540.
XX
PR 25-APR-1994; 94FR-0004937.
XX
PA (ASRE-) ASSOC DEV RECH & GENETIQUE MOLECULAIRE.
XX
PI Egly J, Humbert S, Moncollin V;
XX
DR WPI; 1995-382993/49.
DR P-PSDB; AAR88225.
XX
PT New protein sub-unit(s) of transcription factor BTF2 - useful for
PT treating or diagnosing deficiencies in DNA repair processes

```
XX
PS   Claim 1; Fig 2; 16pp; French.
XX
CC   This is the nucleotide sequence of the transcription factor BTF2 p44
CC   subunit gene.  The sequence encodes a protein of 395 amino acids.
CC   The genes for the p34 (AAT03477) and p44 subunits were isolated from a
CC   HeLa DNA library in lambda-ZAPII using oligonucleotide probes and
CC   primers based on microsequencing of the purified subunits (e.g.
CC   AAT03479-80). Neither the p34 nor the p44 subunits contain any kinase,
CC   ATPase or helicase activity and cannot be used to reconstitute BTF2
CC   activity even with the p62 and p89 BTF2 subunits in an in vitro
CC   transcription system. The proteins can be used to raise antibodies useful
CC   for detecting abnormally low levels of the subunits.  The DNA sequences
CC   can be used similarly for DNA levels. The antibodies and probes are
CC   useful in the detection of development of cancer, partic. skin melanoma
CC   but also xeroderma or Cockayne syndrome, skin pigmentation disorders or
CC   sensitivity to UV light.
XX
SQ   Sequence 1185 BP; 356 A; 220 C; 255 G; 354 T; 0 other;


  Query Match             2.8%;  Score 19;  DB 16;  Length 1185;
  Best Local Similarity   100.0%;  Pred. No. 9.7;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      62 tgaagaacctgaaagaact 80
           ||||||||||||||||||
Db       6 tgaagaacctgaaagaact 24


RESULT   7
AAC78156
ID   AAC78156 standard; cDNA; 1816 BP.
XX
AC   AAC78156;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human cancer associated gene sequence SEQ ID NO:550.
XX
KW   Human; cancer associated gene; cancer antigen; detection; cancer;
KW   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW   antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW   antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW   dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW   vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW   immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW   allergic reaction; graft versus host disease; organ rejection;
KW   haemostatic; thrombolytic; cardiovascular disorder; infection;
KW   neurological disease; drug screening; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055350-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05882.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587533/55.
DR   P-PSDB; AAB43947.
XX
PT   Novel isolated nucleic acids comprising sequences encoding peptides
PT   useful for treating or diagnosing e.g. cancer -
XX
PS   Claim 1; Page 1075-1076; 2352pp; English.
XX
CC   AAC77607 to AAC78448 encode the human cancer associated proteins given
CC   in AAB43398 to AAB44239. The proteins can have activities based on the
CC   tissues and cells the genes are expressed in. Example of activities
CC   include: cytostatic; proliferative; vulnerary; immunomodulator;
CC   antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC   antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC   dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC   nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC   polynucleotides and polypeptides can be used for preventing, treating or
CC   ameliorating medical conditions and diagnosing pathological conditions.
CC   Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC   the present invention may be used to treat immune disorders by activating
CC   or inhibiting the proliferation, differentiation or mobilisation of
CC   immune cells, to treat disorders of haematopoietic cells, autoimmune
CC   disorders, allergic reactions, graft versus host disease and organ
CC   rejection, modulate haemostatic or thrombolytic activity, modulate
CC   inflammation, cancers, cardiovascular disorders, neurological disease and
CC   bacterial or viral infections. The peptides, nucleotides, antibodies,
CC   agonists and antagonists may be also be used in drug screens. AAC78449 to
CC   AAC78457 and AAB44240 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 1816 BP; 561 A; 292 C; 398 G; 563 T; 2 other;


  Query Match             2.8%;  Score 19;  DB 21;  Length 1816;
  Best Local Similarity   100.0%;  Pred. No. 9.8;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      62 tgaagaacctgaaagaact 80
           ||||||||||||||||||
Db     223 tgaagaacctgaaagaact 241


RESULT   8
AAA75580
ID   AAA75580 standard; DNA; 3072 BP.
XX
AC   AAA75580;
XX
DT   22-JAN-2001  (first entry)
XX
DE   DNA encoding a mouse zalpha11 ligand polypeptide.
XX
KW   zalpha11 ligand; cytokine; haematopoietic cell proliferation; lymphoma;
KW   tumourigenesis; leukaemia; hematopoiesis; B cell tumour; ss.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             54..494
FT                   /*tag=  a
FT                   /product= "Zalpha11"
XX
PN   WO200053761-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-US06067.
XX
PR   09-MAR-1999;   99US-0264908.
PR   11-MAR-1999;   99US-0265992.
PR   01-JUL-1999;   99US-0142013.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Novak JE,  Presnell SR,  Sprecher CA,  Foster DC,  Holly RD,  Gross JA;
PI   Johnston JV,  Nelson AJ,  Dillon SR,  Hammond AK;
XX
DR   WPI; 2000-565600/52.
DR   P-PSDB; AAB18624.
XX
PT   New human cytokine, designated zalpha11 ligand, useful for stimulating
```

```
PT   the proliferation and/or development of haematopoietic cells in vitro
PT   and in vivo, and for treating tumourigenesis -
XX
PS   Disclosure; Page 220-222; 256pp; English.
XX
CC   The present sequence encodes a mouse zalpha11 ligand polypeptide,
CC   which is a cytokine. The zalpha11 ligand is useful for stimulating the
CC   proliferation and development of haematopoietic cells in vitro and in
CC   vivo. Zalpha11 ligand polynucleotides can be used as primers or probes
CC   for cloning the zalpha11 gene. The zalpha11 ligand is useful for
CC   treating tumourigenesis. A zalpha11 ligand-saporin fusion toxin may be
CC   used for treating leukaemias and lymphomas. Antagonists against zalpha11
CC   ligand are useful as research reagents for characterizing ligand-receptor
CC   interaction. Antagonists are also useful for inhibiting expansion,
CC   proliferation, activation and differentiation of cells involved in
CC   regulating hematopoiesis. The zalpha11 ligand may also be used to
CC   stimulate an immune response against B cell tumour, a virus, a parasite
CC   or a bacterium. The zalpha11 polypeptides, polynucleotides, antagonists,
CC   agonists and antibodies are also useful for the detection, diagnosis,
CC   prevention, and treatment of diseases associated with a zalpha11 ligand
CC   genetic defect.
XX
SQ   Sequence 3072 BP; 925 A; 591 C; 623 G; 933 T; 0 other;


  Query Match             2.8%;  Score 19;  DB 21;  Length 3072;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    535 tgtgaagatggcaatgaaa 553
          ||||||||||||||||||
Db   1257 tgtgaagatggcaatgaaa 1275


RESULT   9
AAX13947
ID   AAX13947 standard; DNA; 811 BP.
XX
AC   AAX13947;
XX
DT   31-MAR-1999  (first entry)
XX
DE   H. pylori GHPO 1275 gene.
XX
KW   GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
KW   peptic ulcer disease; ss.
XX
OS   Helicobacter pylori.
XX
FH   Key             Location/Qualifiers
FT   CDS             51..764
FT                   /*tag= a
XX
PN   WO9843478-A1.
XX
PD   08-OCT-1998.
XX
PF   01-APR-1998;   98WO-US06371.
XX
PR   29-JUL-1997;   97US-0902615.
PR   01-APR-1997;   97US-0833457.
PR   24-JUN-1997;   97US-0881227.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
XX
PI   Al-Garawi A,  Kleanthous H,  Miller C,  Oomen RP,  Tomb J;
XX
DR   WPI; 1998-542293/46.
DR   P-PSDB; AAW98228.
XX
PT   New isolated Helicobacter polynucleotides - used to develop products
PT   for the diagnosis, prevention and treatment of Helicobacter
PT   infections and gastrointestinal diseases
XX
PS   Claim 1; Page 164-165; 2054pp; English.
XX
CC   This sequence represents a polynucleotide of the invention. It was
CC   isolated from Helicobacter pylori and encodes a H.pylori GHPO protein.
CC   The polypeptides can be used for preventing or treating Helicobacter
CC   infections, and gastroduodenal diseases associated with these
CC   infections, including acute, chronic, and atrophic gastritis, and peptic
CC   ulcer diseases, e.g. gastric and duodenal ulcers. They can also be used
CC   for the production of antibodies. The products can also be used for
CC   detection and diagnosis.
XX
SQ   Sequence 811 BP; 253 A; 146 C; 187 G; 225 T; 0 other;


  Query Match             2.6%;  Score 18;  DB 19;  Length 811;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    480 atccacacctttaaaatt 497
          ||||||||||||||||||
Db  *  729 atccacacctttaaaatt 746


RESULT  10
AAD10125/c
ID   AAD10125 standard; cDNA; 2324 BP.
XX
AC   AAD10125;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Mouse serotransferrin (siderophilin) cDNA.
XX
KW   Mouse; cytostatic; antiinflammatory; immunoregulatory; tissue integrity;
KW   wound healing; immune response; vaccine; cancer; asthma; allergy;
KW   cell trafficking; therapy; secreted protein; serotransferrin;
KW   siderophilin; Tf; beta-1-metal binding globulin; transferrin; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             43..2136
FT                   /*tag= a
FT                   /product= "Mouse serotransferrin (siderophilin)"
XX
PN   WO200148192-A1.
XX
PD   05-JUL-2001.
XX
PF   21-DEC-2000; 2000WO-NZ00256.
XX
PR   23-DEC-1999;   99US-0171678.
PR   28-NOV-2000; 2000US-0724864.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Watson JD,  Murison JG;
XX
DR   WPI; 2001-425665/45.
DR   P-PSDB; AAE05358.
XX
PT   Novel isolated polypeptide useful to isolate corresponding interacting
PT   proteins or other compounds, to quantitatively determine levels of
PT   interacting proteins or other compounds, and as therapeutic target  -
XX
PS   Claim 1; Page 61-62; 101pp; English.
XX
CC   The patent discloses novel polynucleotides and their corresponding
CC   proteins which play a major role in induction of growth, cell migration
CC   and proliferation, cell-cell interaction and the differentiation of
CC   tissue-specific cells. These proteins are important in the maintenance
```

```
CC   of tissue integrity and thus are important in wound healing. They are
CC   useful in various assays to determine the biological activity, to raise
CC   antibodies, to isolate corresponding interacting proteins or other
CC   compounds, to quantitatively determine levels of interacting proteins or
CC   other compounds, and as therapeutic target in a whole range of disease
CC   states. Compositions comprising the novel proteins of the invention are
CC   useful for treating mammalian disorders. Polynucleotides of the invention
CC   are useful in genome and physical mapping, in positional cloning of
CC   genes, to tag or identify an organism or its reproductive material (as
CC   non-disruptive tags for marking organisms), and for the diagnosis and
CC   treatment of mammalian diseases which is the consequence of inappropriate
CC   expression of kinase genes. They are useful for promoting immune response
CC   as part of a vaccine or anti-cancer treatment, as target for cancer
CC   treatment, as immunoregulatory and anti-inflammatory molecule, as
CC   diagnostic for specific types of cancer and for development of an
CC   anti-cancer treatment, and as a target for antagonists in the treatment
CC   of diseases such as asthma and allergy. They are also useful to inhibit
CC   or enhance the activity of the soluble molecule that binds proteins of
CC   the invention, for tissue and neural regeneration, to promote or block
CC   cell trafficking, and as anti-inflammatory and/or vaccine adjuvant.
CC   The present sequence is a cDNA encoding mouse serotransferrin
CC   (siderophilin). Serotransferrin (Tf) also known as beta-1-metal
CC   binding globulin is a part of the transferrin family.
XX
SQ   Sequence 2324 BP; 592 A; 585 C; 627 G; 520 T; 0 other;


  Query Match              2.6%;  Score 18;  DB 22;  Length 2324;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    274 tcgaatttgcagtggttg 291
          ||||||||||||||||||
Db   1538 TCGAATTTGCAGTGGTTG 1521


RESULT  11
AAT44520/c
ID   AAT44520 standard; DNA; 4118 BP.
XX
AC   AAT44520;
XX
DT   24-FEB-1997  (first entry)
XX
DE   NTHI hxuC + hxuB gene region.
XX
KW   HxuC; HxuB; NTHI; vaccine; genetic immunisation; diagnosis;
KW   meningitis; pneumonia; bacteraemia; otitis media; ss.
XX
OS   Haemophilus influenzae nontypeable strain N182.
XX
FH   Key             Location/Qualifiers
FT   CDS             142..2301
FT                   /*tag= a
FT                   /product= HxuC
FT   CDS             2376..4073
FT                   /*tag= b
FT                   /product= HxuB
XX
PN   WO9633275-A1.
XX
PD   24-OCT-1996.
XX
PF   15-APR-1996;   96WO-US05167.
XX
PR   20-APR-1995;   95US-0425843.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Cope LD,  Hansen EJ,  Hanson MS,  Jarosik GP;
XX
DR   WPI; 1996-485781/48.
DR   P-PSDB; AAW01462, AAW01465.
XX
PT   Genes encoding H. influenzae HxuC and HxuB surface-expressed
PT   protein(s) - useful in the prepn. of vaccines for children against
PT   H. influenzae infection
XX
PS   Claim 12; Page 142-147; 188pp; English.
XX
CC   A PCR fragment (AAT44520) obtd. from nontypeable Haemophilus
CC   influenzae (NTHI) strain N182 includes 2 open reading frames that
CC   respectively encode surface-expressed HxuC (AAW01462) and HxuB
CC   (AAW01465) proteins involved in haem regulation.  PCR was performed
CC   on genomic DNA using the primers given in AAT44522-23.  A similar
CC   DNA fragment (AAT44519) coding for NTHI HxuC (AAW01461) and HxuB
CC   (AAW01464) was obtd. from strain TN106.  The genes can be used: to
CC   produce recombinant HxuC and HxuB proteins for use in vaccines; to
CC   design probes used to diagnose NTHI infection; and in genetic
CC   immunisation to protect against NTHI infection.
XX
SQ   Sequence 4118 BP; 1393 A; 720 C; 795 G; 1210 T; 0 other;


  Query Match              2.6%;  Score 18;  DB 17;  Length 4118;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    233 caaacaagttttccagtg 250
          ||||||||||||||||||
Db   2221 CAAACAAGTTTTCCAGTG 2204


RESULT  12
AAX13097
ID   AAX13097 standard; DNA; 6253 BP.
XX
AC   AAX13097;
XX
DT   19-MAR-1999  (first entry)
XX
DE   Enterococcus faecalis genome contig SEQ ID NO:160.
XX
KW   Enterococcus faecalis; contig; detection; Enterococcal infection;
KW   vaccine; attenuation; computer readable medium; ds.
XX
OS   Enterococcus faecalis.
XX
PN   WO9850555-A2.
XX
PD   12-NOV-1998.
XX
PF   04-MAY-1998;   98WO-US08985.
XX
PR   14-NOV-1997;   97US-0066009.
PR   06-MAY-1997;   97US-0044031.
PR   16-MAY-1997;   97US-0046655.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Barash SC,  Dillon PJ,  Kunsch CA;
XX
DR   WPI; 1999-045171/04.
XX
PT   New isolated Enterococcus faecalis polynucleotides and polypeptides
PT   - used to develop products for the detection of Enterococcus and for
PT   use in vaccines for prevention or attenuation of Enterococcus
PT   infection.
XX
PS   Claim 1; Page 902-905; 2084pp; English.
XX
CC   A computer readable medium has been developed which has recorded on it
CC   982 nucleotide sequences isolated from the Enterococcus faecalis genome.
CC   AAX12938 to AAX13919 represent these nucleotide sequences which are
CC   primary nucleotide sequences, also known as contigs. The computer-based
CC   system can identify fragments of the Enterococcus faecalis genome with
```

```
CC   commercial importance. The products can be used to detect the presence
CC   of Enterococcus faecalis in samples. They can also be used for
CC   diagnosing Enterococcal infection in an animal and monitoring
CC   progression of disease, and for identifying agents which can be used to
CC   modulate the growth or pathogenicity of Enterococcus faecalis, or
CC   another related organism, in vivo or in vitro. In particular the
CC   polypeptides encoded by the Enterococcus faecalis nucleotide sequences
CC   can be used in vaccines to prevent or attenuate an Enterococcal
CC   infection.
XX
SQ   Sequence 6253 BP; 1966 A; 1261 C; 1046 G; 1963 T; 17 other;


  Query Match             2.6%;  Score 18;  DB 20;  Length 6253;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     536 gtgaagatggcaatgaaa 553
           ||||||||||||||||||
Db    4908 gtgaagatggcaatgaaa 4925


RESULT  13
AAZ01425/c
ID   AAZ01425 standard; DNA; 1038602 BP.
XX
AC   AAZ01425;
XX
DT   07-OCT-1999  (first entry)
XX
DE   Complete genome sequence of Chlamydia trachomatis.
XX
KW   Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW   paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;
KW   nongonococcal uretritis; epidymitis; cervicitis; salpingitis;
KW   bartholinitis; pneumopathy; venereal lymphogranulomatosis; ss.
XX
OS   Chlamydia trachomatis.
XX
PN   WO9928475-A2.
XX
PD   10-JUN-1999.
XX
PF   27-NOV-1998;   98WO-IB01939.
XX
PR   04-NOV-1998;   98US-0107077.
PR   28-NOV-1997;   97FR-0015041.
PR   17-DEC-1997;   97FR-0016034.
XX
PA   (GEST ) GENSET.
XX
PI   Griffais R;
XX
DR   WPI; 1999-371125/31.
XX
PT   Genome sequence of Chlamydia trachomatis
XX
PS   Claim 1; Page 373-656; 1755pp; English.
XX
CC   The present sequence represents the complete genome of Chlamydia
CC   trachomatis. Open reading frames (ORFs) of the genome encode
CC   polypeptides AAY36754-Y37949. The polypeptides can be used as vaccines
CC   against Chlamydia trachomatis. Antisense and ribozyme sequences can also
CC   be used to control growth of the microorganism. Chlamydia trachomatis is
CC   responsible for a large number of diseases, e.g. eye diseases such as
CC   conventional trachoma, nonendemic trachoma, paratrachoma, and inclusion
CC   conjunctivitis; genital diseases such as nongonococcal uretritis,
CC   epidymitis, cervicitis, salpingitis, perihepatitis, bartholinitis;
CC   pneumopathy in breast feeding infants; and venereal
CC   lymphogranulomatosis. The polypeptides of the invention may be of use in
CC   treating these diseases.
XX
SQ   Sequence 1038602 BP; 304265 A; 214645 C; 214259 G; 305001 T; 432 other;


  Query Match             2.6%;  Score 18;  DB 20;  Length 1038602;
  Best Local Similarity   100.0%;  Pred. No. 34;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     451 gtggagctcccaactgcc 468
           ||||||||||||||||||
Db  964676 GTGGAGCTCCCAACTGCC 964659


RESULT  14
AAI19482/c
ID   AAI19482 standard; DNA; 156 BP.
XX
AC   AAI19482;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #9415 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 9415; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 156 BP; 37 A; 43 C; 51 G; 25 T; 0 other;


  Query Match             2.5%;  Score 17;  DB 22;  Length 156;
  Best Local Similarity   100.0%;  Pred. No. 96;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      33 ctcacatcttccagcag 49
           |||||||||||||||||
Db     137 CTCACATCTTCCAGCAG 121
```

```
RESULT  15
AAI20472/c
ID   AAI20472 standard; DNA; 156 BP.
XX
AC   AAI20472;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #10405 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 10405; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 156 BP; 37 A; 43 C; 51 G; 25 T; 0 other;


  Query Match             2.5%;  Score 17;  DB 22;  Length 156;
  Best Local Similarity   100.0%;  Pred. No. 96;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      33 ctcacatcttccagcag 49
           |||||||||||||||||
Db     137 CTCACATCTTCCAGCAG 121



Search completed: February 26, 2002, 13:25:05
Job time: 5821 sec
```